DEVELOPMENTAL BIOLOGY **183,** 21–36 (1997)
ARTICLE NO. DB968485

GATA-5: A Transcriptional Activator Expressed in a Novel Temporally and Spatially-Restricted Pattern during Embryonic Development

Edward E. Morrisey, Hon S. Ip, Zhihua Tang, Min M. Lu, and Michael S. Parmacek

Department of Medicine, University of Chicago, Chicago, Illinois 60637

Members of the GATA family of zinc finger transcription factors regulate critical steps of cellular differentiation during vertebrate development. In the studies described in this report, we have isolated and functionally characterized the murine GATA-5 cDNA and protein and defined the temporal and spatial pattern of GATA-5 gene expression during mammalian development. The amino terminus of the mouse GATA-5 protein shares high level amino acid sequence identity with the murine GATA-4 and -6 proteins, but not with other members of the GATA family. GATA-5 binds to the functionally important CEF-1 nuclear protein binding site in the cardiac-specific slow/cardiac troponin C (cTnC) transcriptional enhancer and overexpression of GATA-5 transactivates the cTnC enhancer in noncardiac muscle cell lines. During embryonic and postnatal development, the pattern of GATA-5 gene expression differs significantly from that of other GATA family members. In the primitive streak embryo, GATA-5 mRNA is detectable in the precardiac mesoderm. Within the embryonic heart, the GATA-5 gene is expressed within the atrial and ventricular chambers (ED 9.5), becomes restricted to the atrial endocardium (ED 12.5), and is subsequently not expressed in the heart during late fetal and postnatal development. Moreover, coincident with the earliest steps in lung development, only the GATA-5 gene is expressed within the pulmonary mesenchyme. Finally, the GATA-5 gene is expressed in tissue-restricted subsets of smooth muscle cells (SMCs), including bronchial SMCs and SMCs in the bladder wall. These data are consistent with a model in which GATA-5 performs a unique temporally and spatially restricted function in the embryonic heart and lung. Moreover, these data suggest that GATA-5 may play an important role in the transcriptional program(s) that underlies smooth muscle cell diversity.

INTRODUCTION

The GATA family of zinc finger transcription factors play important roles in regulating cell lineage differentiation during vertebrate development (for review see (Orkin, 1992; Simon, 1995; Weiss and Orkin, 1995a). Six GATA family members have been identified in vertebrate species. These six proteins have been classified into two subfamilies based on common amino acid sequence identities and patterns of expression (Arceci *et al.*, 1993; Dorfman *et al.*, 1992; Evans *et al.*, 1988; Heikinheimo *et al.*, 1994; Ho *et al.*, 1991; Jiang and Evans, 1996; Joulin *et al.*, 1991; Kelley *et al.*, 1993; Ko *et al.*, 1991; Laverriere *et al.*, 1994; Lee *et al.*, 1990; Morrisey *et al.*, 1996; Tsai *et al.*, 1989; Wilson *et al.*, 1990; Yamamoto *et al.*, 1990; Zon *et al.*, 1991). Each GATA family member contains a conserved C-X_2-C-X_{17}-C-X_2-C (type IV) zinc finger DNA-binding domain that recognizes a consensus sequence motif (WGATAR) located within transcriptional regulatory elements that control sets of lineage-specific genes (Ho *et al.*, 1991; Ko and Engel, 1993; Merika and Orkin, 1993; Omichinski *et al.*, 1993; Yang *et al.*, 1994). In addition to activating transcription directly, recent studies have revealed that GATA-1 may directly, or indirectly, be involved in regulating programmed cell death during development (Weiss and Orkin, 1995b; Blobel and Orkin, 1996). Gene targeting experiments have revealed that while the GATA-1, -2, and -3 proteins are expressed in overlapping patterns in the hematopoietic cell lineages, mice harboring null mutations in each of these genes exhibit distinct phenotypes (Pandolfi *et al.*, 1995; Pevny *et al.*, 1991; Simon *et al.*, 1992; Tsai *et al.*, 1994; Weiss *et al.*, 1994). These data are consistent with a model wherein each GATA family member regulates unique lineage-restricted developmental programs.

0012-1606/97 $25.00

Previous studies have suggested that the GATA-4/5/6 subfamily of transcription factors may direct tissue-specific gene expression during development of the vertebrate heart (Jiang and Evans, 1996; Kelley *et al.*, 1993; Laverriere *et al.*, 1994; Tamura *et al.*, 1993). Functionally important GATA binding sites have been identified within the promoters of multiple cardiac-specific genes, including slow/cardiac troponin C, α-myosin heavy chain, B-type natriuretic peptide (BNP), and cardiac troponin I, and each of these cardiac-specific transcriptional regulatory elements can be transactivated by overexpression of GATA-4 (and in some cases GATA-6) in noncardiac muscle cell lines (Grepin *et al.*, 1994; Ip *et al.*, 1994; Molkentin *et al.*, 1994; Morrisey *et al.*, 1996). In addition, expression of GATA-4 antisense transcripts in pluripotent P19 embryonal carcinoma cells blocks the development of beating cardiac myocytes and interferes with expression of cardiac-specific contractile protein isoforms (Grepin *et al.*, 1995). Moreover, ectopic expression of *Xenopus* GATA-4, -5, or -6 during embryogenesis is capable of prematurely activating expression of the genes encoding α-cardiac actin and α-myosin heavy chain (Jiang and Evans, 1996).

In avian and *Xenopus* species the genes encoding GATA-4, -5, and -6 are expressed in an overlapping pattern in the heart and gut epithelium (Jiang and Evans, 1996; Laverriere *et al.*, 1994). Thus, it remains unclear whether each of these GATA subfamily members subserves unique or redundant developmental functions. In this regard, it is noteworthy that in mammalian species, significant differences in the developmentally regulated, tissue-restricted patterns of GATA-4 and -6 gene expression have been identified that are not present, or were not appreciated, in lower vertebrates (Heikinheimo *et al.*, 1994; Morrisey *et al.*, 1996). For example, coincident with the onset of vasculogenesis, the murine GATA-6 gene is expressed in arterial and venous smooth muscle cells (SMCs), while the GATA-4 gene is not (Morrisey *et al.*, 1996). Moreover, a mammalian GATA-5 homologue has not, as yet, been described. However, in avian species, the GATA-5 gene is expressed at least as early as stage 7 within the heart and the gene continues to be expressed at high levels within both the endocardium and myocardium throughout postnatal development (Laverriere *et al.*, 1994). Taken together, these data suggest that although the GATA-4, -5 and -6 genes are expressed in an overlapping pattern during embryogenesis and postnatal development, fine differences in the spatial and temporal pattern of expression of each family member may impact profoundly on their developmental function.

In the studies described in this report, we have isolated the murine GATA-5 cDNA and compared the structure of the deduced murine GATA-5 protein to each of the previously characterized GATA family members. We have demonstrated that forced expression of GATA-5 transactivates a cardiac-specific transcriptional regulatory element in noncardiac muscle cell lines. In addition, the mouse GATA-5 cDNA was used as a molecular probe to map the temporal and spatial patterns of GATA-5 gene expression in the embryonic mouse and during postnatal development. These data are consistent with a model wherein GATA-5 serves as a transcriptional activator that regulates a novel developmental program within the embryonic heart and lung. In addition, these data suggest that GATA-5 may play an important role in the transcriptional program(s) that underlies smooth muscle cell diversity.

MATERIALS AND METHODS

Isolation of Murine GATA-5 cDNA Clones

To isolate the murine GATA-5 cDNA, 10^6 recombinant clones from a Day 13 mouse embryonic heart λZAP cDNA library (Stratagene) were screened under moderate stringency conditions (0.5X SSC, 0.1% SDS at 50°C) with a radiolabeled murine GATA-4 cDNA probe encoding the two zinc fingers and adjacent basic domain (bp 1173–1547) as described previously (Parmacek and Leiden, 1989). Replicate filters were hybridized to murine GATA-4 3′-untranslated (3′UTR) and GATA-6 3′UTR cDNA probes in order to identify recombinant clones encoding each of these previously characterized cDNAs (Arceci *et al.*, 1993; Morrisey *et al.*, 1996) and these clones were not characterized further. Two positively hybridizing clones (λmG5-A and λmG5-B), which were approximately 1.8- and 3.2-kb in size, respectively, were isolated, purified to homogeneity, and sequenced with an ABI 377 Automated DNA Sequencer using the Sequencher computer program (Gene Codes, Ann Arbor, MI). The deduced murine GATA-4, -5, and -6 amino acid sequences were aligned using the MAP Multiple Sequence Alignment computer program (Huang, 1994).

Northern Blot Analyses

Northern blot analyses were performed as described previously (Solway *et al.*, 1995). Two nonoverlapping GATA-5 cDNA probes were utilized in these analyses. The mouse GATA-5 ORF cDNA probe (bp 253–707) encodes 52-bp of the 5′UTR and 403-bp of the GATA-5 open reading frame (ORF); while the mouse GATA-5 3′UTR cDNA (bp 2262–2740) is derived exclusively from the 3′UTR (see Fig. 1). Quantitative image analyses were performed using a Molecular Dynamics PhosphorImager (Sunnyvale, CA).

Plasmids

The murine GATA-5 expression plasmid, pcDNAG5, was subcloned by digesting the λmG5-A cDNA clone with *Eco*RI and *Kpn*I to isolate the 1828-bp insert (see Fig. 1A) and subcloning the resultant cDNA fragment encoding the 404-aa GATA-5 protein into the *Eco*RV site of the eukaryotic expression plasmid, pcDNA3 (Invitrogen). The murine GATA-4 expression plasmid, pcDNAG4, has been previously described (Morrisey *et al.*, 1996). The p-124cTnCGH and p-124cTnCGHμCEF-1 growth hormone (GH) reporter plasmids contain the 124-bp cTnC cardiac-specific cTnC promoter-enhancer and the 124-bp cTnC promoter-enhancer with a five nucleotide mutation within the CEF-1 nuclear protein binding site, respectively, subcloned into the p0GH human GH reporter plasmid (Nichol's Institute) and have been described previously

(Morrisey *et al.*, 1996). The pCEF-1GH and pμCEF-1GH expression plasmids contain a single copy of the murine cTnC CEF-1 nuclear protein binding site or a copy of the cTnC CEF-1 nuclear protein binding site containing a mutation in the GATA motif, respectively, subcloned immediately 5′ of the minimal rabbit β-globin promoter in the p0GH plasmid and have been described previously (Morrisey *et al.*, 1996).

Cell Culture and Transfections

NIH 3T3 and COS-7 cells were grown as described previously (Parmacek *et al.*, 1994; Parmacek and Leiden, 1989; Solway *et al.*, 1995) For transient cotransfection experiments, 10^6 NIH 3T3 cells were cotransfected with 25 μg of the appropriate GATA expression plasmid, 1 μg of pMSVβgal reference plasmid, 2.5 μg of the appropriate GH reporter plasmid, and 20 μl of Lipofectin Reagent (Gibco). Forty-eight hours following transfection, cell supernatants were assayed for GH by radioimmunoassay (Nichol's Institute) and cell lysates were prepared and analyzed for β-galactosidase activity and protein content as described previously (Ip *et al.*, 1994). Data are expressed as normalized growth hormone ± SEM. All experiments were repeated at least three times to ensure reproducibility.

Electrophoretic Mobility Shift Assays (EMSAs)

Nuclear extracts were prepared from COS-7 cells 48 hr following transient transfection with the pcDNAG5, pcDNAG4, or pcDNA3 expression plasmids as described (Ip *et al.*, 1994). The following complementary oligonucleotides were synthesized with *Bam*HI/*Bgl*II overhanging ends:

cTnC CEF-1 5′ CCAGCCT**GAGATTA**CA**GGGAG** 3′
cTnC μCEF-1 5′ CCAGCCTG**GGGCCC**CAGGGAG 3′
αMHC-GATA 5′ CCAAAGGGCCGATGGGC**AGATAG**AGGA-GAGACAGGA 3′

The native and mutated GATA sites are shown in bold font. EMSAs were performed as described previously (Ip *et al.*, 1994). Antibody supershifts were performed by preincubating nuclear extract with 1 μl of goat polyclonal α-GATA-4 affinity purified IgG (that was raised against an epitope corresponding to amino acids 420–439 of the mouse GATA-4 protein (Santa Cruz, Biotech., Cat. No. sc-1237)) for 20 min at 25°C prior to performing the binding reaction as described previously (Ip *et al.*, 1994). This antiserum does not cross-react with the mouse GATA-1, -2, or -3 proteins.

In Situ Hybridization of Staged Murine Embryos

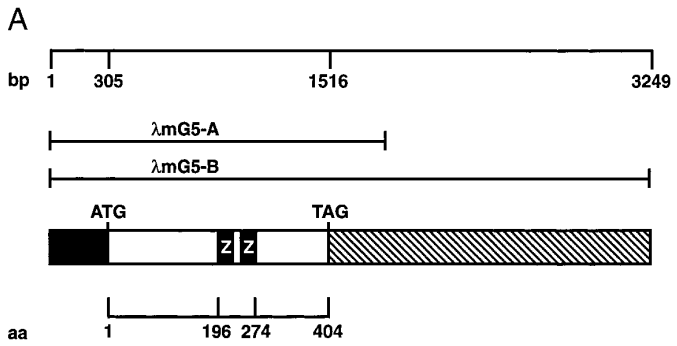
In situ hybridization was performed essentially as described by Eichele and coworkers (Kuratani *et al.*, 1994). The murine GATA-5 (bp 253–707), GATA-4 (bp 486–936), GATA-6 (bp 589–1119), and SM22α cDNAs (bp 333–1102) were subcloned into pGEM7Z or pGEM11Z and were *in vitro*-transcribed using T7 or SP6 polymerase in the presence of ^{35}S-labeled UTP to generate sense and antisense cRNA probes, respectively. Sections were processed for emulsion autoradiography and were poststained with Hoecht 33258 and visualized by epifluorescence and darkfield microscopy on a Zeiss Axiophot microscope. To detect nonspecific background, hybridizations were performed with the radiolabeled sense riboprobe to alternate sections under identical conditions.

RESULTS

Isolation and Structural Characterization of Murine GATA-5 cDNA Clones

To isolate the murine GATA-5 cDNA, 10^6 recombinant clones from an Embryonic Day (ED) 13 mouse heart cDNA library were screened under moderate stringency conditions with a subfragment of the murine GATA-4 cDNA encoding the evolutionarily conserved zinc finger domains. In addition, replicate filters were hybridized to murine GATA-4 3′-untranslated (3′UTR) and GATA-6 3′UTR cDNA probes in order to identify recombinant clones encoding each of these previously characterized cDNAs (Arceci *et al.*, 1993; Morrisey *et al.*, 1996). Two plaques that hybridized to the mouse GATA-4 zinc finger cDNA probe, but not the murine GATA-4 or GATA-6 3′UTR cDNA probes, designated λmG5-A and λmG5-B, were identified using this screening strategy. The λmG5-A and λmG5-B clones contained overlapping cDNA inserts of approximately 1.8- and 3.2-kb, respectively (Fig. 1A). Each clone contained a 1212-bp open reading frame (ORF) that encoded a predicted 404-amino acid polypeptide with a calculated molecular mass of 42.1 kDa (Fig. 1A). The initiation codon conforms to the consensus sequence described by Kozak (1984). The 1.4-kb difference in the length of the two cDNA clones resulted from differences in the length of their respective 3′UTRs (Fig. 1A). Notably, neither cDNA contained a consensus polyadenylation signal.

Comparison of the deduced amino acid sequence of the murine GATA-5 protein with the previously isolated chicken and *Xenopus* GATA-5 proteins (Jiang and Evans, 1996; Laverriere *et al.*, 1994) revealed 62 and 53% amino acid sequence identity, respectively. In contrast, comparison of the deduced amino acid sequence of the murine GATA-5 protein with the chicken GATA-4 and -6 proteins (Laverriere *et al.*, 1994) revealed 44 and 48% amino acid sequence identity. Despite the relatively low level of amino acid sequence identity between the *Xenopus*, chicken, and murine GATA-5 proteins, multiple amino acid residues were identified that are identical in the chicken, *Xenopus*, and murine GATA-5 proteins that are not conserved in the GATA-4 and -6 proteins (data not shown); suggesting that these cDNAs encode the murine homologue of the GATA-5 protein. Of note, comparison of the deduced amino acid sequence of the murine GATA-5 protein with the previously isolated murine GATA-4 and GATA-6 proteins (Arceci *et al.*, 1993; Morrisey *et al.*, 1996) did reveal high level amino acid sequence identity across the two zinc finger domains and adjacent C-terminal basic domain (Fig. 2, shaded gray). In addition, discrete regions located within the amino terminus of the mouse GATA-5 protein, that



B

```
                                                                                                                                ACAACCCACCGGTCACCCTCCCTGGGCCAACGCCGCGCCTTCC
TGTGAGGGCTTTGGCTGGCGATCTCTGCGCAAGGCAAGGACCTCGTTCCTAGGGTCCTGCGCCTCATCTAACCCTGTCATTCCACCC
AGAGGAGCCCGGGACCTGCCCTGGTGAACCGACTCAAAAAAGTGGAGTGCCCAAGCGGAGCAGCAGCACCTGCCCCGGGCGAGCGCA
AGAGCTGATGGAGGGGGCCCAGTGATCCTGCCCTAGACGTCTGCCCGTCTTTCTTCTCTGCAGGTCAAGCTCGCGCGGGGAAAAAAA

ATG TAC CAA AGC TTG GCG TTA GCC CAG AGC CCC GGC CAG GGC ACT TAC GCC GAC TCC GGA GCC TTT
 M   Y   Q   S   L   A   L   A   Q   S   P   G   Q   G   T   Y   A   D   S   G   A   F
CTG CAC TCT TCG GGA ACC GGC TCC CCG GTG TTC GTG GCA CCC ACT CGC ATG CCC TCC ATG CTG CCG
 L   H   S   S   G   T   G   S   P   V   F   V   A   P   T   R   M   P   S   M   L   P
TAC CTG CCT TCG TGT GAG CCG GGC TCG CAG GCT CCC GCG CTC GCC GCG CAC TCC AGC TGG ACA CAA
 Y   L   P   S   C   E   P   G   S   Q   A   P   A   L   A   A   H   S   S   W   T   Q
ACA GTT GCC GCT GAT TCC TCA GCC TTC GGC TCC GGA AGC CCG CAC CCG CCA GCC GCA CAC CCA CCG
 T   V   A   A   D   S   S   A   F   G   S   G   S   P   H   P   P   A   A   H   P   P
GGA GCC ACC ACC TTC CCG TTC GCT CAC AGC CCT CCG GGG TCG GGC AGC GGC GGC AGC GCG GGG GTC
 G   A   T   T   F   P   F   A   H   S   P   P   G   S   G   S   G   G   S   A   G   V
CGG GAC GGC GGC GCC TTC CAG GGC GCG CTG CTG GCT CGC GAA CAG TAC CCG ACC CCG CTC GGG AGA
 R   D   G   G   A   F   Q   G   A   L   L   A   R   E   Q   Y   P   T   P   L   G   R
CCC ATG GGC GCC TCG TAC CCC ACA ACC TAC CCA GCA TAC ATG AGT TCC GAC GTA GCC CCT TCG TGG
 P   M   G   A   S   Y   P   T   T   Y   P   A   Y   M   S   S   D   V   A   P   S   W
ACC TCC GGA GCC TTC GAC AGC AGC ATC CTG CAC GGT CTG CAA GCT CGC CCT GGT GGC CTC CCT GGC
 T   S   G   A   F   D   S   S   I   L   H   G   L   Q   A   R   P   G   G   L   P   G
CGC AGG ACC AGC TTC GTA CCT GAC TTC TTG GAG GAG TTC CCT GGT GAG GGC CGG GAG TGC GTC AAC
 R   R   T   S   F   V   P   D   F   L   E   E   F   P   G   E   G   R   E   C   V   N
TGT GGA GCC CTG TCT ACA CCA CTG TGG CGC AGG GAC GGC ACC GGA CAC TAT CTA TGC AAT GCC TGC
 C   G   A   L   S   T   P   L   W   R   R   D   G   T   G   H   Y   L   C   N   A   C
GGC CTC TAT CAC AAG ATG AAC GGG GTC AAC CGA CCG CTA GTG AGG CCC CAA AAG CGT CTG TCC TCA
 G   L   Y   H   K   M   N   G   V   N   R   P   L   V   R   P   Q   K   R   L   S   S
TCC CGA AGA TCC GGC CTA TGC TGC TCC AAC TGC CAT ACT GCC ACC ACC ACC CTC TGG AGG CGC AAC
 S   R   R   S   G   L   C   C   S   N   C   H   T   A   T   T   T   L   W   R   R   N
TCG GAG GGC GAA CCA GTG TGC AAC GCC TGT GGC CTC TAC ATG AAG CTC CAT GGG GTG CCA AGG CCA
 S   E   G   E   P   V   C   N   A   C   G   L   Y   M   K   L   H   G   V   P   R   P
CTG GCA ATG AAA AAG GAG AGT ATC CAG ACA AGA AAA CGG AAG CCA GAG AAC CCT GCC AAA ATC AAG
 L   A   M   K   K   E   S   I   Q   T   R   K   R   K   P   E   N   P   A   K   I   K
GGC TCC TCA GGA TCC ACA GCA AAC ACC ACA GCC TCT TCT CCC ACT CTC CTC AAC TCT GAG AGC TCA
 G   S   S   G   S   T   A   N   T   T   A   S   S   P   T   L   L   N   S   E   S   S
GCC ACC ACT TTG AAG GCA GAG TCC AGT CTG GCA TCT CCA GTG TGT GCT GGG CCC ACT ATC ACC TCC
 A   T   T   L   K   A   E   S   S   L   A   S   P   V   C   A   G   P   T   I   T   S
CAG GCC TCC AGC CCA GCG GAC GAA TCC CTG GCC TCC AGC CAT TTG GAG TTC AAG TTC GAA CCT GAA
 Q   A   S   S   P   A   D   E   S   L   A   S   S   H   L   E   F   K   F   E   P   E
GAC TTT GCC TTC ACC TCC TCG TCC ATG AGC CCC CAG GCT GGT CTC AGT GGG GTC CTA CGT CAG GAG
 D   F   A   F   T   S   S   S   M   S   P   Q   A   G   L   S   G   V   L   R   Q   E
ACC TGG TGT GCT CTG GCC TTG GCC TAG GTGCTCAGGAAACTGTCACCACAGTACATCCCAGGAACAGAAGTTAGTTCC
 T   W   C   A   L   A   L   A
TGTGTCAGCAGCCCCAGAAGTCCCTCACCACTATGCTGGCGAAATTTTCTGTGCTGATTGCCTGGAGAAAGAACAGGCTGGTGGATG
AAGGGTTGGAGCCTCAGCCGGGGTAGTAATGGACTCCCTAGACCCCTCATGGACCTCTCTGCAGCCAAGATGACTGGAGATATCCAC
CAATTACTTGAATTTAGCTCTTGACATCGATCGTCAAAATTTACAGTTTGTAGCAGCTGGCAAAGGTTCATAGTATCCCACAGTAAA
ATCTGATCCAAGTGAAACAAACATACTCCAGGCAGGCAATAAGAAATAGGGCCTCTGGGCCCTCGTCGTCTTCATTTCCCGTGTTTG
ACTCTGTACGACTGAGAGAGCACACTCTTGCTTACAGACGCAGGACTGCCAGAGACCAGTGTACAGGTTGAGCACTGTACCAGGACA
CCGCCCCAAGGCAAACGCATCCCTCCGGCCCATCATCGCAGACTTCTGTATTTATTTGTATGTTTGTTTCCAGGCCAGGCAGTAAAG
GGCCTTGTTCTGGCAATCTGCGTACTAAGGTGTTTGCTAGTGGGCAGACCTGCAGAGACACTGTTGGAATTGGGTGATGTGTGACCT
TGTGGCCTAGCTCAGCCATTATTCCAGCAGTTGGCAACACAGCCCTATTTGTTAGCATCTTTTTCACCCAGGTCACATCAGCTGGTG
GTGACTTTGCCCAAGGCCTGCAGGTCAGTTCTAGGGCCCTGGTACAGGGAGGTGACAGGCAACAAAAGACAGGTGTTCTATATGCCG
CATTCCCACAGTTTTAGCTCAACGGGTAATTCCTGGCCATCCCAGGTTTCCATTCCCCAGTGAGGAAACCAAGAAGCAAGGGAGTCA
GGGATCATCCACCCATGAAGGTCACCCCTAACCTGTCTCCTGCGTGGACCCTCAATCCTTCCTTCCAGTTCAGCTTGAGGGAGAGGC
AGTAGGCCCTGGTTAGAAGGAAATCAAGAGGCAGTAGCCTGCAGCATTTGCAGGCATCGCAGGGGCAAAGGCAAGACATGCTCTTGT
AACTAGCTCCAAGCAGGCATACCTCACCACCCGTGTATCTGTGGAAGAAAGCCAGCTGCCTAGGTGTCTGTTATCACATCCCAGAGG
CTGGAGTAAGGAGAGGAGGGGACTAGAGACCTGGACACAGGCAACACCGTGTGAGCATCTCCAAGCACCCATACCCCAAGTGTGCAC
AGATGCAGACAGGACCACAGCTGTGCACAAGGTCATAGCACCATGTTCTCGATACAGTGCTGTTCCTAGGAGCTCTGTGTGTTCCTA
GTGTCCTAGCCTCAGTTTCCCCTTCTGTGATTAGAGAGGTGTCTGTGAGCACATCCGTAATTCATGTACCTTTGAGGTGGTTCAAGG
CACTCACCAACCTAGGAAGGCTTCCTGGAGACAGCCACAGTGTTAAGGCACAAGGCAGATGGCACCCCCTTGAATGGAGAGGAAGGG
AACTCGGGAAGAGAGGGGCATGAAGACATAAGATCCCTTCCACGTTGGTTTCTCTGCCTTTCTTTGTTTCATTCGTTTGTTTTTTGT
TGTTGTTATTTTGTTTTGTTTTGTTTTTTGAGACAGGGCCTTACTATGTAGCTCTGGCTGTCCTGGAACTCATTATGTAGACCAGGC
TGGCCTCAAACTCACAGAGACCTGCC
```

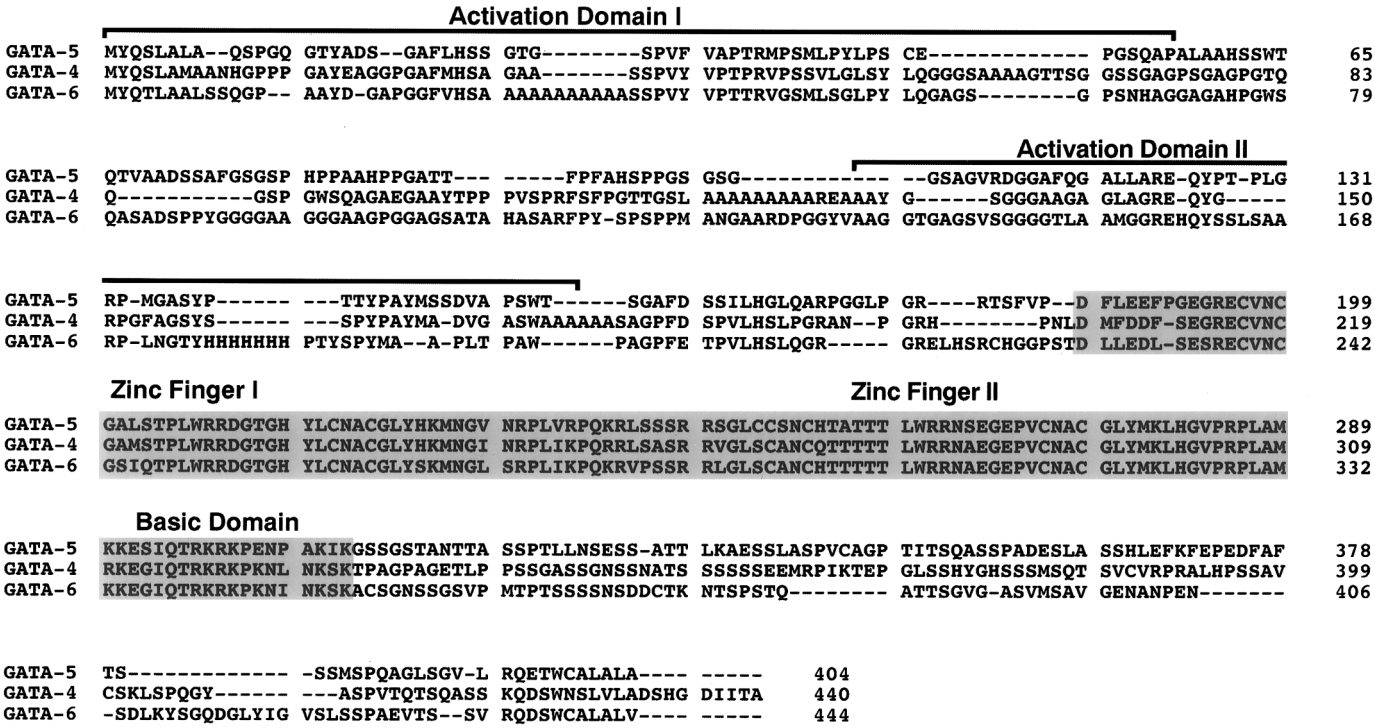


FIG. 2. Comparison of the amino acid sequences of the murine GATA-4, GATA-5, and GATA-6 proteins. The protein sequences were aligned using the MAP multiple protein sequence alignment algorithm (Huang, 1994). The conserved zinc fingers and basic domain are shown in gray. Activations Domains I and II of the mouse GATA-4 protein (Morrisey *et al.*, submitted) are shown above the amino acid sequences. Amino acid residue numbers are shown to the right of each sequence.

have been demonstrated previously to encode independent transcriptional activation domains in the mouse GATA-4 protein (Morrisey *et al.*, submitted), were also conserved in the predicted mouse GATA-5 protein (Fig. 2). In contrast, the amino terminus of the mouse GATA-5 protein does not share identifiable sequence homology with either the GATA-1, -2, or -3 proteins (data not shown). Outside of the zinc fingers, C-terminal basic domain and short regions within the amino terminus of the murine GATA-5 protein, the amino acid sequences of the GATA-4, -5, and -6 proteins have diverged significantly (Fig. 2).

Tissue-Specificity of GATA-5 Gene Expression

To determine the tissue-specificity of GATA-5 gene expression during postnatal development, a murine GATA-5 cDNA probe, derived from the 5′UTR and open reading frame (ORF) (bp 253–707), was radiolabeled and used to probe Northern blots containing RNA prepared from adult murine tissues. As shown in Fig. 3 (upper panel), the GATA-5 ORF cDNA probe hybridized to three transcripts approximately 1.8, 3.2, and 8.0 kb (arrows) in size. GATA-5 mRNA was detected in the small intestine, stomach, and bladder and at very low levels within the lung (arrows, Fig. 3, upper panel, lanes 3, 4, 7, and 8). In the stomach and small intestine the 3.2-kb transcript was most abundant, while in the bladder the 1.8- and 8.0-kb transcripts predominated. Even upon prolonged autoradiographic exposure GATA-5 mRNA was not detected within the adult heart (Fig. 3, upper panel, lane 1). In addition, GATA-5 mRNA was not detected in RNA prepared from the adult aorta, skeletal muscle, kidney, liver, spleen, thymus, uterus, and brain (Fig. 3, upper panel,

FIG. 1. The structure and deduced amino acid sequence of the mouse GATA-5 cDNA and protein encoded by clones λmG5-A and λmG5-B. (A) A schematic representation of the structure of the λmG5-A and λmG5-B cDNA clones. The size of the two cDNA clones (λmG5-A and λmG5-B) encoding the murine GATA-5 protein is shown above the map. The 5′-untranslated region is shaded. The protein coding region is shown as an open box with the two evolutionarily conserved zinc finger domains indicated (Z). The 3′-untranslated region is hatched. The size of the deduced mouse GATA-5 protein in amino acids (aa) is shown below the map. (B) The nucleotide sequence and deduced amino acid sequence of the murine GATA-5 protein encoded by λmG5-A and λmG5-B. The deduced amino acid sequence of the murine GATA-5 protein is shown below the nucleotide sequence.

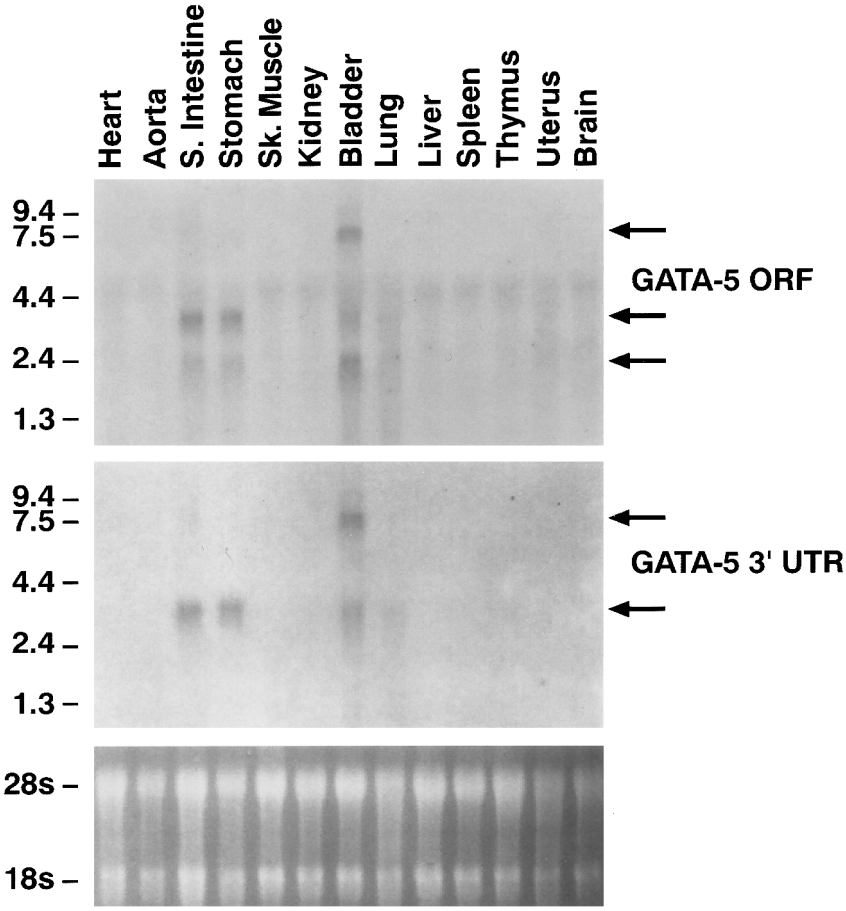


FIG. 3. The *in vivo* tissue distribution of GATA-5 gene expression during postnatal development. The top panel shows a Northern blot analysis of RNA samples isolated from adult murine tissues hybridized to the radiolabeled GATA-5 cDNA probe (bp 253–707) which includes 52 bp of the 5′ UTR and 403 bp of the GATA-5 open reading frame (GATA-5 ORF). This GATA-5 cDNA probe hybridized to 1.8-, 3.2-, and 8.0-kb transcripts (arrows), which were present in the small intestine, stomach, bladder, and lung (at very low levels). RNA size markers are shown in kilobases to the left of the blot. The middle panel shows the same Northern blot hybridized to the radiolabeled GATA-5 3′-untranslated region cDNA probe (bp 2262–2740) (GATA-5 3′ UTR). This GATA-5 cDNA probe hybridized exclusively to the 3.2- and 8.0-kb transcripts (arrows). The bottom panel shows the ethidium bromide-stained gel prior to membrane transfer of RNA. The locations of the 28S and 18S ribosomal RNA bands are indicated to the left of the gel.

lanes 2, 5, 6, 9–13). Of note, the 1.8- and 3.2-kb GATA-5 transcripts that were identified by Northern blot analyses corresponded in size to the λmG5-A and λmG5-B cDNA clones, respectively. In order to determine which, if any, of the three GATA-5 transcripts contain the 1.4 kb of distal 3′ sequence that was present only in the λmG5-B cDNA clone (see Fig. 1A), the Northern blot analysis was repeated using a mouse GATA-5 cDNA probe derived from the distal 3′UTR (bp 2262–2740) (Fig. 3, middle panel). In contrast to the GATA-5 ORF cDNA probe that hybridized to all three GATA-5 transcripts (upper panel), the 3′UTR probe hybridized exclusively to the 3.2- and 8.0-kb transcripts (arrows, Fig. 3, middle panel). Taken together, during postnatal development GATA-5 gene expression is restricted to the gut, bladder, and lung (at low levels). In this regard, the tissue-restricted pattern of expression of the GATA-5 gene during postnatal development differs markedly from that of the GATA-4 which is expressed in the heart, gut, testes, and ovary, and GATA-6 which is expressed in the heart, gut, bladder, and vasculature (Heikinheimo *et al.*, 1994; Morrisey *et al.*, 1996).

The GATA-5 Gene Has a Unique Lineage-Restricted Pattern of Expression during Embryonic Mammalian Development

To determine the temporal and spatial pattern of GATA-5 gene expression during mammalian development, a series

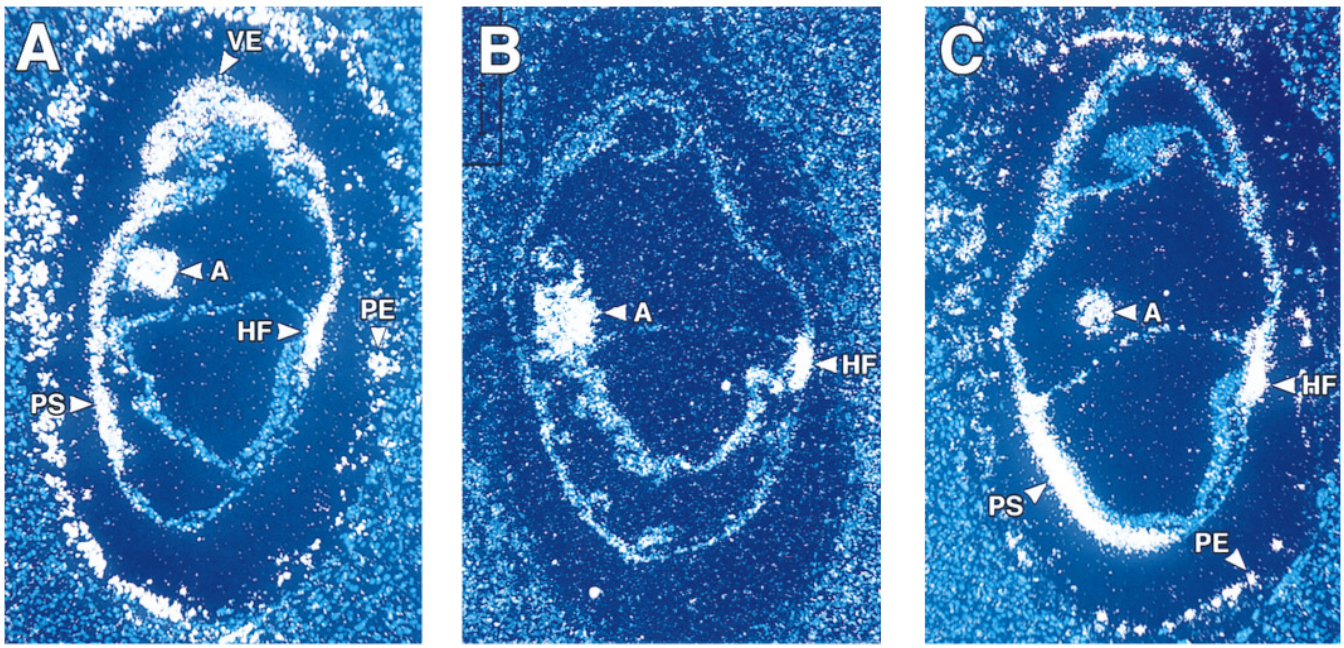


FIG. 4. Expression of the GATA-4, -5, and -6 genes in the primitive streak embryo. *In situ* hybridization analyses were performed using radiolabeled GATA-4 (A), GATA-5 (B), and GATA-6 (C) antisense riboprobes on sections of ED 7.0-staged murine embryos. (A) Specific hybridization (white) of the GATA-4 antisense riboprobe to the embryonic mesoderm in the primitive streak (PS) region and mesoderm subjacent to the headfold (HF) was visualized. Extra-embryonic hybridization of the GATA-4 riboprobe to the allantois (A), visceral endoderm (VE), and parietal endoderm (PE) was also observed. Original magnification, 12.5X. (B) Specific hybridization of the GATA-5 antisense riboprobe to the embryonic mesoderm subjacent to the headfold region (HF) containing the cardiogenic plate and the allantois (A) was visualized. (C) Specific hybridization of the GATA-6 antisense riboprobe to the embryonic mesoderm in the primitive streak (PS) region and mesoderm subjacent to the headfold (HF) was observed. Extra-embryonic hybridization of the GATA-6 riboprobe to the allantois (A) and parietal endoderm (PE) was also observed.

of *in situ* hybridization experiments were performed on staged murine embryos using GATA-4-, -5-, and -6-specific riboprobes. Within the primitive streak embryo (ED 7.0), the GATA-5 riboprobe hybridized exclusively to the embryonic mesoderm subjacent to the future headfold (HF) region which contains the cardiogenic plate (Fig. 4B). In contrast, to this restricted pattern of expression, the GATA-4 and -6 riboprobes hybridized to the lateral plate mesoderm of the primitive streak (PS) region, as well as, the mesoderm subjacent to the headfold (HF) region (Figs. 4A and 4C). Previous studies have suggested that GATA-4 may play an important role in formation of extra-embryonic membranes (Soudais *et al.*, 1995). Thus, it is noteworthy that in contrast to the GATA-4 gene which was expressed at high levels within the visceral endoderm (VE) and the GATA-4 and -6 genes which were expressed within the parietal endoderm (PE) (Figs. 4A and 4C and (Morrisey *et al.*, 1996)), GATA-5 mRNA was only observed within the allantois (A) (Fig. 4B). In summary, in primitive streak embryos, the GATA-5 gene was expressed in a more restricted pattern than either the GATA-4 or -6 genes in both the precardiac mesoderm and within extra-embryonic tissues.

In the ED 9.5 embryo, GATA-5 mRNA was detected within both the common atria (A) and ventricle (V) as well as within the cardiac outflow tract or truncus arteriosus (TA) (Fig. 5A). High power epifluorescence exposures revealed that the GATA-5 riboprobe hybridized more intensely to the common atrial chamber (A) than the common ventricle (V) and most intensely to the underlying septum transversum (Fig. 6A). In contrast, the GATA-4 and -6 riboprobes hybridized at equal intensity to the common atrial and ventricular chambers (data not shown and (Morrisey *et al.*, 1996)). By midgestation (ED 12.5), the spatial pattern of expression of the mouse GATA-5 gene within the heart (Figs. 5B and 6B) diverged significantly from that of the GATA-4 and -6 genes. Coincident with cardiac septation, GATA-5 mRNA was detected exclusively within the atria and was restricted primarily to endocardial cells (white arrowheads) lining the atria and endocardial cushions (EC) (Fig. 6B). In contrast, both the GATA-4 and -6 genes were expressed throughout the atrial and ventricular myocardium and endocardium at this developmental stage (Morrisey *et al.*, 1996). Moreover, by ED 16.5, GATA-5 mRNA was no longer detected within the heart (Fig. 5C). Of note,

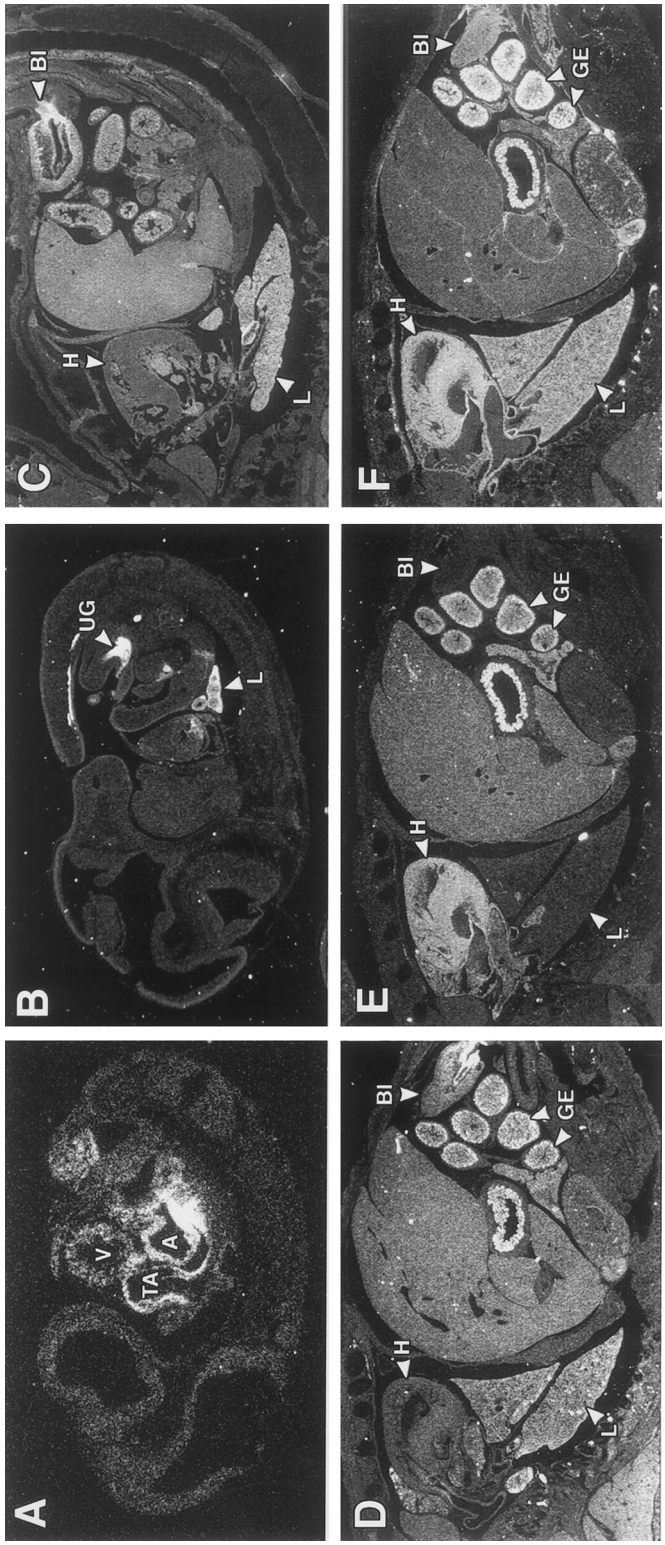


FIG. 5. The temporal and spatial pattern of GATA-5 gene expression during embryonic development of the mouse. *In situ* hybridization analyses were performed on a staged series of murine embryos using the GATA-5 antisense riboprobe (A–D) or the GATA-4 (E) or GATA-6 (F) antisense riboprobes as described under Materials and Methods. (A) Hybridization of the GATA-5 antisense riboprobe to an ED 9.5 embryo. The GATA-5 riboprobe hybridized (white signal) to the common atria (A), ventricle (V) (at low levels), and truncus arteriosus (TA) (see also Fig. 6A). In addition, hybridization of the GATA-5 riboprobe to the cells lining the primitive mid- and hindgut was also observed (not shown). Original magnification, 5X. (B) Hybridization of the GATA-5 antisense riboprobe to an ED 12.5 embryo. The GATA-5 riboprobe hybridized to the lung bud (L) (see also Fig. 7A), urogenital ridge (UG) (see also Fig. 6C), gut epithelium (see also Fig. 6E), and at low levels to the atrial endocardium (see also Fig. 6B). Original magnification, 3.1X. (C) Hybridization of the GATA-5 antisense riboprobe to an ED 16.5 embryo. The GATA-5 riboprobe hybridized to the bladder wall (B1), gut epithelium, and lung (L). Of note, the GATA-5 riboprobe did not hybridize to the heart (H) at this stage of embryonic development. Original magnification, 3.1X. (D) Hybridization of the GATA-5 riboprobe to an ED 18.5 embryo. The GATA-5 riboprobe hybridized to the bladder wall (B1), intestinal (GE) and stomach epithelium, and lung parenchyma (L). Original magnification, 3.1X. (E and F) Hybridization of the GATA-4 antisense riboprobe (E) and the GATA-6 antisense riboprobe (F) to ED 18.5 embryos. In contrast to the pattern of GATA-5 gene expression in the late fetal embryo, the GATA-4 (E) and the GATA-6 (F) riboprobes hybridized to the heart (H). The GATA-4 and -6 genes were also expressed in the gut epithelium (GE). Only GATA-6 mRNA was detected in the bladder wall (B1), the lung (L) (see also Figs. 7E and 7H), and the arteries and veins. Original magnification, 3.1X.

the low grade signal observed in the atria is artefactual and emanates from trapped RBCs along the atrial wall. In contrast to the absence of GATA-5 gene expression within the heart at ED 18.5 (Fig. 5D), the GATA-4 (Fig. 5E) and GATA-6 (Fig. 5F) genes were expressed at high levels throughout the atrial and ventricular myocardium and endocardium during late fetal development. Thus, within the heart, the GATA-5 gene has a temporally and spatially restricted pattern of expression which is distinct from that of both the GATA-4 and -6 genes during embryonic and postnatal development.

At midgestation (ED 12.5), the GATA-5 gene is expressed most abundantly within the lung bud (L) and urogenital sinus (UG) (Fig. 5B). Within the nascent lung bud at ED 12.5, GATA-5 gene expression was restricted to pulmonary mesenchyme (Fig. 7A). In contrast, at this developmental stage (ED 12.5), the GATA-6 gene was expressed within epithelial cells lining the bronchi and bronchioles (Fig. 7B, Br), while the GATA-4 gene was not expressed within the lung (data not shown). The pattern of GATA-5 gene expression within the lung parenchyma overlaps with, but is much more diffuse than, that of the smooth muscle cell marker, SM22α gene, which was expressed exclusively within bronchial (Br) and pulmonary arterial SMCs (Fig. 7C). In the fetal lung at ED 18.5, the GATA-5 riboprobe hybridized intensely to the smooth muscle cells surrounding the large airways (Figs. 7D and 7G, open arrowheads (Br)), while low level hybridization was observed throughout the lung parenchyma. Conversely, GATA-6 gene expression was detected within pulmonary arterial SMCs (Figs. 7E and 7H, white arrowheads (A)), and at low levels throughout the lung parenchyma (Fig. 7H). In contrast to the arterial SMC-specific or bronchial SMC-specific pattern of GATA-5 and -6 gene expression, respectively, within the fetal lung the SM22α gene was expressed in both bronchial (Br) and pulmonary arterial (A) SMCs (Figs. 7F and 7I). These data demonstrate that the GATA-5 and -6 genes are expressed in distinct, complementary, lineage-restricted patterns in the embryonic lung.

During postnatal development, GATA-5 gene expression was detected within the gastrointestinal tract and bladder (see Fig. 3). During embryogenesis, intense hybridization of the GATA-5 riboprobe to the dorsal aspect of the urogenital ridge can be appreciated at least as early as ED 12.5 (Figs. 5B and 6C, white arrowheads). Over the next 48 hr, the pattern of hybridization expands circumferentially to surround the epithelial cells lining the urogenital sinus (Fig. 6D). Subsequently by ED 18.5, the GATA-5 gene is expressed at high levels within the wall of the embryonic bladder (B1) (Fig. 5D) where the gene continues to be expressed throughout postnatal development. Within the developing gastrointestinal tract, the GATA-5 riboprobe hybridized to cells lining the embryonic mid- and hindgut at least as early as ED 9.5 (data not shown). Of note, despite the fact that the GATA-5 gene is expressed in bronchial SMCs (Figs. 7D and 7G) and SMCs of the bladder wall (Figs. 6D and 7D), within the gut GATA-5 gene expression is restricted to the epithelial cells (Fig. 6E, white arrowheads). This epithelial-specific pattern of expression within the gastrointestinal tract continues as the intestinal crypts and villi differentiate during the fetal period (ED 18.5) (Fig. 6F). These data demonstrate that the GATA-5 gene is an early developmental marker of the gut epithelium and of bladder smooth muscle cells where the gene continues to be expressed throughout postnatal development.

GATA-5 Binds to the Cardiac-Specific cTnC Transcriptional Enhancer

GATA-4 and -6 have been demonstrated to bind to the cardiac-specific cTnC transcriptional enhancer and transactivate the cTnC gene in noncardiac muscle cell lineages (Ip *et al.*, 1994; Morrisey *et al.*, 1996). The developmentally regulated, cell- and chamber-specific pattern of GATA-5 gene expression suggested the hypothesis that GATA-5 might recognize a unique set of transcriptional targets. Therefore, to determine whether GATA-5 binds to the functionally important nonconsensus CEF-1/GATA motif located within the cTnC promoter-enhancer (Ip *et al.*, 1994), EMSAs were performed using nuclear extracts prepared from COS-7 cells transiently transfected with the GATA-5 expression plasmid, pcDNAG5. As shown in Fig. 8, the radiolabeled cTnC CEF-1 oligonucleotide probe bound two nuclear protein complexes (arrows) that were not identified in nuclear extracts prepared from COS-7 cells transfected with the pcDNA3 control expression plasmid (Fig. 8, lanes 1 and 2). Each of these complexes was abolished by the addition of unlabeled cTnC CEF-competitor oligonucleotide or by an unlabeled oligonucleotide corresponding to the functionally important consensus GATA motif identified within the α-cardiac myosin heavy chain promoter (Molkentin *et al.*, 1994) (Fig. 8, lanes 3 and 4), but not by nonspecific oligonucleotide competitor (Fig. 8, lane 5). Of note, neither nuclear protein complex was supershifted following preincubation with α-GATA-4-specific IgG (Fig. 8, lane 6). In addition, three nuclear protein complexes were identified in binding reactions containing nuclear extracts prepared from COS-7 cells transfected with the mouse GATA-4 expression plasmid, pcDNAG4, and the radiolabeled CEF-1 probe (Fig. 8, lane 7). Each of these nuclear protein complexes were ablated and supershifted (dashed arrows) following preincubation with α-GATA-4-specific IgG (Fig. 8, lane 8). These data demonstrate that GATA-5 can bind specifically to DNA, and that GATA-5 binds to the functionally important nuclear protein binding site in the cTnC cardiac-specific promoter-enhancer.

GATA-5 Transactivates the cTnC Cardiac-Specific Transcriptional Enhancer in Noncardiac Muscle Cells

To determine whether forced expression of GATA-5 can transactivate the cardiac-specific cTnC promoter-enhancer

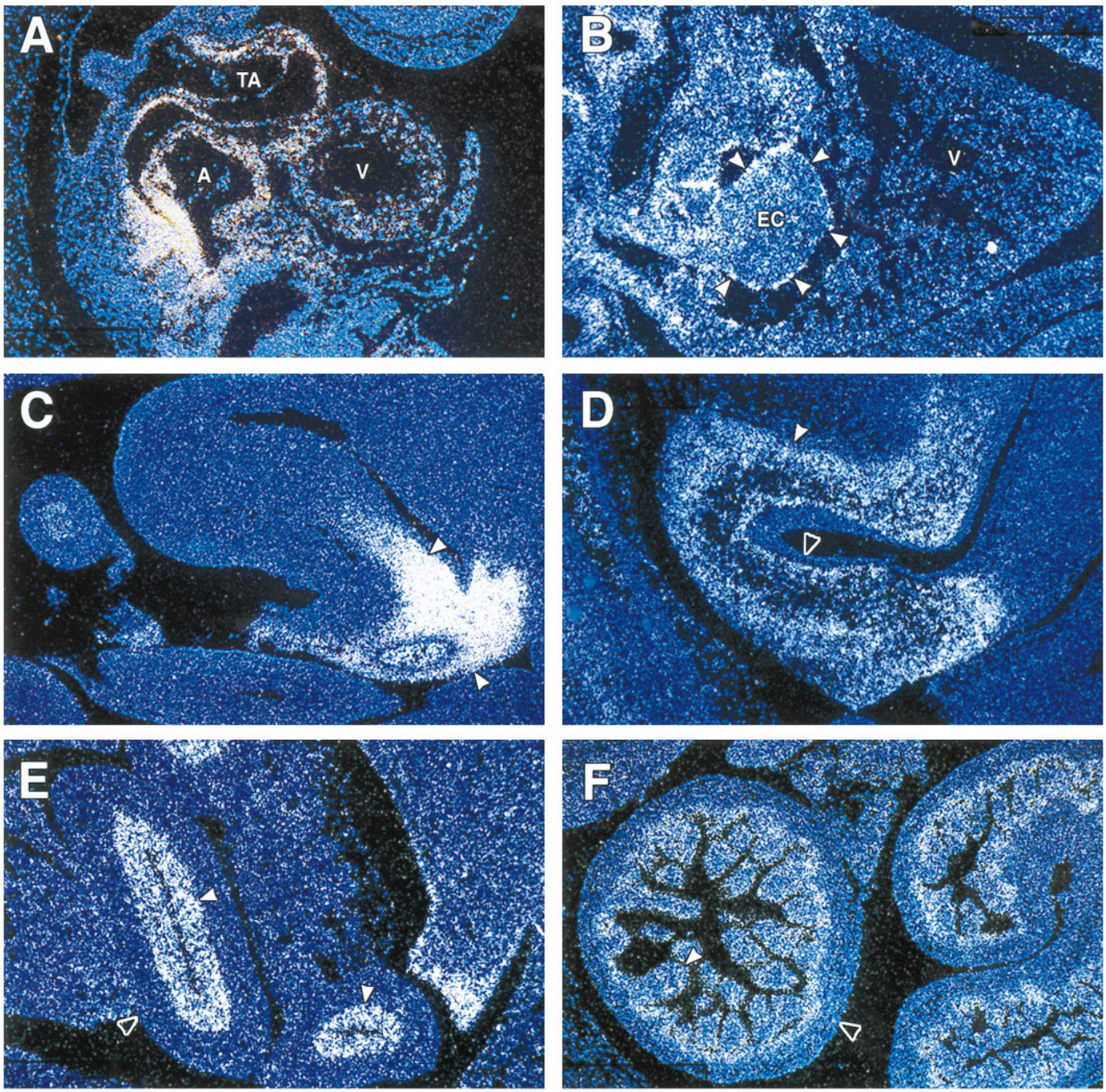
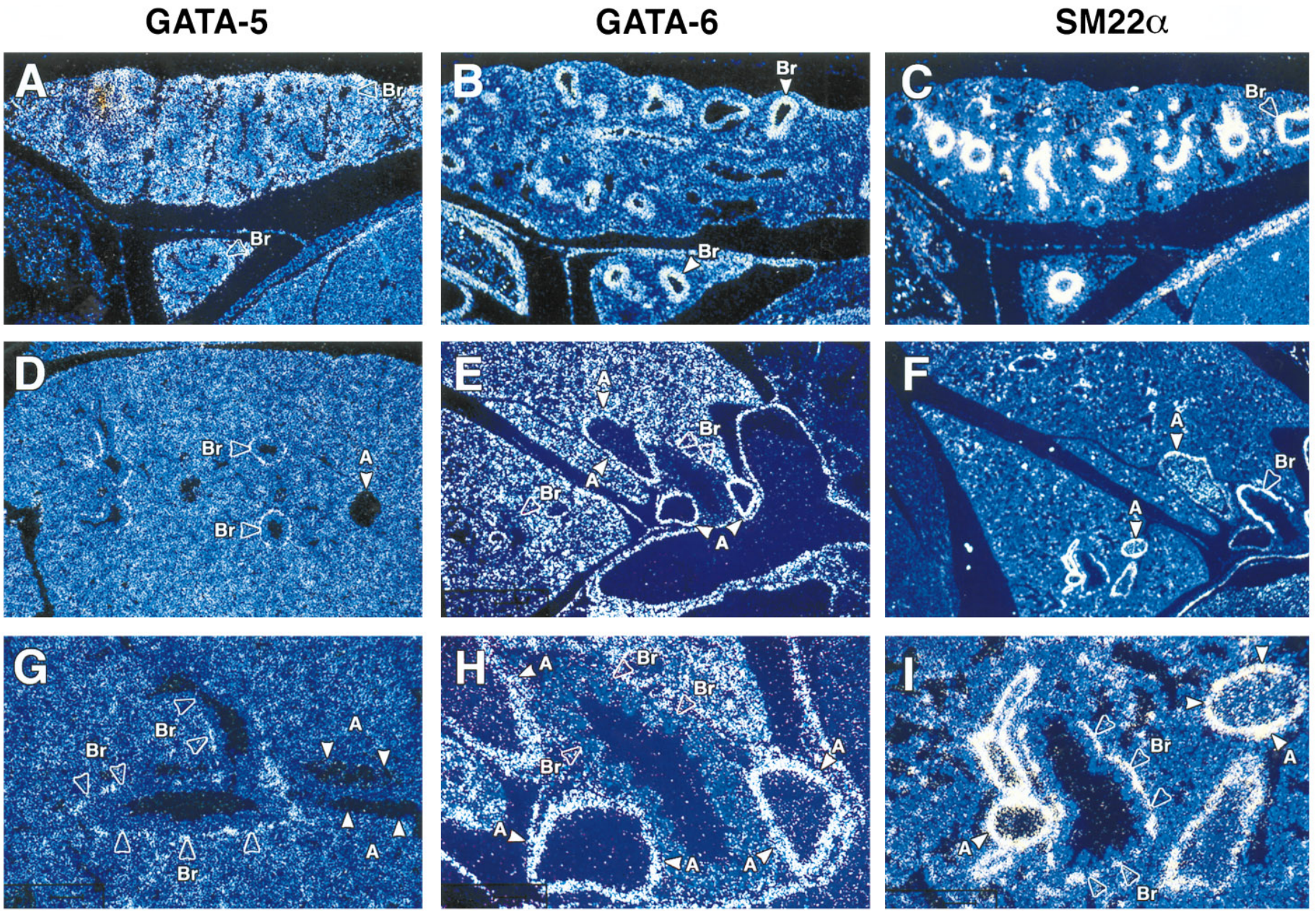


FIG. 6. The GATA-5 gene is expressed in the embryonic heart, the urogenital ridge, and the gut epithelium in a developmentally regulated fashion. *In situ* hybridization analyses were performed using a GATA-5 antisense riboprobe and sections through staged ED 9.5 (A), 12.5 (B, C and E), 14.5 (D), and 18.5 embryos (F). (A) Within the primitive heart (ED 9.5), abundant hybridization (white signal) to the atrial (A) myocardium and endocardium and truncus arteriosus (TA) was visualized. In contrast, only low level hybridization to the ventricular chamber (V) was observed. Original magnification, 10X. (B) At ED 12.5, hybridization to the atrial endocardium (arrowheads) and endocardium lining the endocardial cushion (EC) was visualized. In contrast, the GATA-5 riboprobe did not hybridize to the ventricular myocardium or endocardium. Original magnification, 10X. (C) At ED 12.5, intense hybridization to the dorsal aspect of the urogenital ridge was observed (arrowheads). Original magnification, 10X. (D) At ED 14.5, GATA-5 mRNA can be detected in cells of the urogenital ridge (white arrowhead) surrounding the epithelial cells lining the urogenital sinus (open arrowhead). Original magnification, 10X. (E) At ED 12.5, hybridization to the epithelial cells (white arrowhead) lining the midgut is visualized. Hybridization to the lamina propria was not observed (open arrowhead). Original magnification, 10X. (F) At ED 18.5, hybridization to the epithelial cells (white arrowhead) lining the crypts and villi of the intestine, but not the lamina propria (open arrowhead), was observed. Original magnification, 5X.


GATA-5
GATA-6
SM22α
A
B
C
D
E
F
G
H
I
Br
A

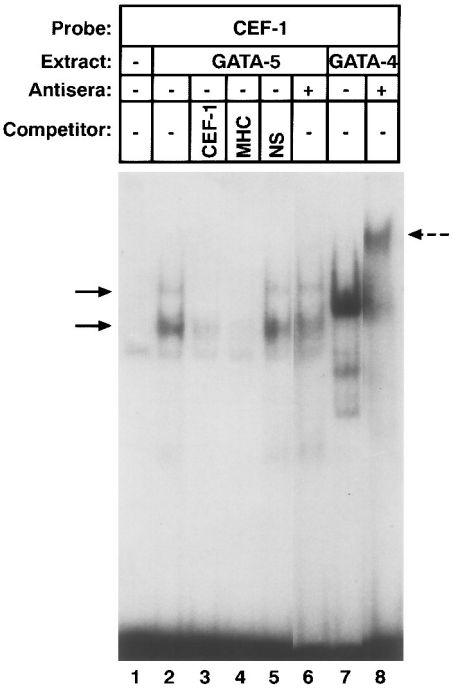


FIG. 8. Binding of GATA-5 and GATA-4 to the CEF-1 site of the cTnC transcriptional enhancer. Radiolabeled CEF-1 oligonucleotides were used in EMSAs with nuclear extracts prepared from COS-7 cells transiently transfected with either the negative control expression plasmid, pcDNA3 (lane 1), or expression plasmids encoding the murine GATA-5 (lanes 2–6) or GATA-4 (lanes 7–8) proteins. Where indicated, binding reactions were preincubated with either 100 ng of indicated unlabeled competitor oligonucleotides or with goat polyclonal α-GATA-4 IgG (+). The bands corresponding to the GATA-5 nuclear protein complexes are shown by arrows to the left of the autoradiogram. The GATA-4 antibody-supershifted complex is indicated with a dashed arrow to the right of the autoradiogram.

in noncardiac muscle cells, NIH 3T3 cells were transiently cotransfected with a GH reporter plasmid driven by the 124-bp cardiac-specific cTnC promoter-enhancer (Parmacek *et al.*, 1992) and an eukaryotic expression plasmid encoding the murine GATA-5 protein. As demonstrated previously (Ip *et al.*, 1994), the negative control plasmid, pcDNA3, failed to transactivate the p-124GH reporter plasmid in NIH 3T3 cells (Fig. 9A, column 1). In contrast, forced expression of GATA-5 resulted in approximately a 275-fold increase in GH reporter activity (Fig. 9A, column 2). To determine whether this increase in transcriptional activity was dependent upon binding of GATA-5 to the CEF-1 motif within the cTnC transcriptional enhancer, the GATA-5 expression plasmid was transiently cotransfected with the p-124GHμCEF-1 reporter plasmid, that contains a mutation within the CEF-1/GATA element that abolishes binding of GATA-5 (data not shown). Mutation of the CEF-1/GATA motif resulted in approximately a 60% reduction in normalized GH activity (Fig. 9A, column 4), demonstrating that most, but not all, of the transcriptional activation of the cTnC promoter was dependent upon an intact CEF-1/GATA motif.

The finding that overexpression of GATA-5 transactivated the p-124GHμCEF-1 plasmid above levels obtained with the negative control expression plasmid, pcDNA3 (Fig. 9A, columns 3 and 4), suggested that GATA-5 might be capable of indirectly activating the cTnC cardiac-specific promoter enhancer. In this regard it is noteworthy that several recent studies have demonstrated that GATA family members can activate transcription either by binding directly to DNA, or alternatively, by binding directly to other transcription factors (Crossley *et al.*, 1995; Osada *et al.*, 1995). Therefore, to determine whether the ability of GATA-5 to transactivate the p-124GHμCEF-1 reporter plasmid was dependent upon any of the other previously characterized *cis*-acting regulatory elements within the cTnC promoter-enhancer (Parmacek *et al.*, 1992), the GATA-5 expression plasmid was transiently cotransfected with the pCEF-1GH reporter plasmid, containing a single copy of the CEF-1 motif upstream of the minimal rabbit β-globin promoter (Morrisey *et al.*, 1996), and the pμCEF-1GH reporter plasmid, containing a mutated copy of the CEF-1 motif subcloned upstream of the minimal β-globin promoter, into NIH 3T3 cells. Forced expression of GATA-5 in NIH 3T3 cells transactivated the pCEF-1GH reporter

FIG. 7. The temporal and spatial patterns of GATA-5, GATA-6, and SM22α gene expression in the murine embryonic and fetal lung. *In situ* hybridization analyses were performed using GATA-5 (A, D, and G), GATA-6 (B, E, and H), and SM22α (C, F, and I) antisense riboprobes on staged ED 12.5 (A–C) and ED 18.5 (D–I) embryonic lung sections. (A–C) At ED 12.5, the GATA-5 riboprobe (A) hybridized diffusely to the pulmonary mesenchyme, but not to the bronchial epithelial cells (Br). In contrast, the GATA-6 riboprobe (B) hybridized to the bronchial epithelial cells (Br), but not to the pulmonary mesenchyme. The SM22α riboprobe (C), a marker of the smooth muscle cell lineage, hybridized to bronchial (Br) and pulmonary arterial SMCs. Original magnification, 10X. (D and G) At ED 18.5, the GATA-5 riboprobe hybridized to the bronchial SMCs (Br, open arrowheads), but not the pulmonary arterial SMCs (A, white arrowheads). In addition, diffuse low level hybridization of the GATA-5 riboprobe to the lung parenchyma was visualized. Original magnification, 5X and 20X. (E and H) At ED 18.5, the GATA-6 riboprobe hybridized to pulmonary arterial SMCs (A, white arrowheads) and to the lung parenchyma. GATA-6 gene expression was not observed in the epithelial cells lining the large airways (Br, open arrowheads). Original magnification, 5X and 20X. (F and I) At ED 18.5, the SM22α riboprobe hybridized exclusively to bronchial (Br) and pulmonary arterial (A) SMCs. Original magnification, 5X and 20X.

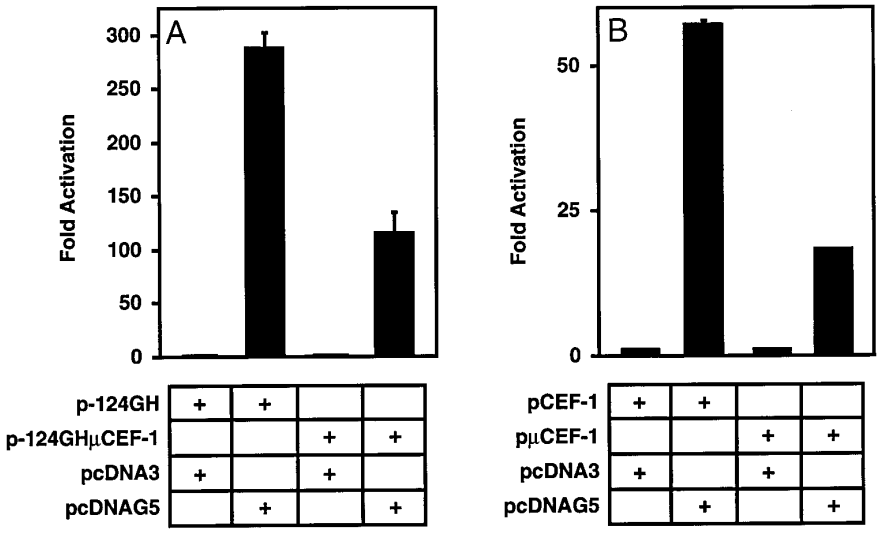


FIG. 9. GATA-5-mediated activation of the cTnC promoter and a minimal CEF-1-containing promoter in noncardiac muscle cells. (A) NIH 3T3 cells were transfected with either 2.5 μg of the p-124cTnCGH (p124GH +) or the p-124cTnCGHμCEF-1 (p-124GHμCEF-1 +) reporter plasmids and either 25 μg of the control expression plasmid (pcDNA3 +) or the GATA-5 expression plasmid (pcDNAG5 +). All transfections contained 1 μg of the pMSVβgal reference plasmid. Forty-eight hours following transfection, the amount of growth hormone in the media and the β-galactosidase activities were determined. Relative GH activity was determined by correcting for differences in transfection efficiencies and normalizing to the amount of growth hormone obtained following transfection of the p-124cTnCGH plasmid with the pcDNA3 expression plasmid. The data are presented as relative GH activity ± SEM. (B) Transient cotransfections were performed as described above with the exception that the pCEF-1GH or pμCEF-1GH reporter plasmids were utilized which contain a single CEF-1 nuclear protein binding site or a mutant CEF-1 nuclear protein binding site, respectively, subcloned immediately 5′ of the minimal rabbit β-globin promoter.

construct 55-fold above levels in cells cotransfected with the negative control expression plasmid, pcDNA3 (Fig. 9B, columns 1 and 2). In contrast, only an 18-fold induction in normalized GH activity was demonstrated with the pμCEF-1GH plasmid (Fig. 9B, columns 2 and 3). Taken together, these data demonstrate that GATA-5 is a potent transcriptional activator that is capable of activating the cardiac-specific cTnC promoter enhancer in noncardiac muscle cells. The majority of the transcriptional activation is dependent upon an intact GATA motif within the cTnC CEF-1 nuclear protein binding site and the residual induction in transcriptional activity results from the indirect activation of the cTnC CEF-1 element.

DISCUSSION

Members of the GATA family of transcription factors play important roles in transducing nuclear events that regulate cell lineage differentiation during vertebrate development. In this report, we have isolated and structurally and functionally characterized the mouse GATA-5 protein. In addition, we have determined the temporal and spatial patterns of GATA-5 gene expression during mouse embryonic development. Structural comparison of the mouse GATA-5 protein with the previously characterized mouse GATA-4 and -6 proteins identified specific regions within two N-terminal activation domains that have been conserved across the GATA-4/5/6 subfamily (Morrisey *et al.*, submitted). In addition, GATA-5 binds to the functionally important CEF-1 nuclear protein binding site in the cardiac-specific cTnC transcriptional enhancer and over-expression of GATA-5 is sufficient to transactivate the cTnC promoter-enhancer in noncardiac muscle cells. In contrast to the GATA-4 and -6 genes, which are developmentally coexpressed in the embryonic and adult heart, the GATA-5 gene is expressed in a temporally and spatially restricted pattern during early embryonic cardiac development. Moreover, coincident with expansion of the embryonic lung bud, the GATA-5 gene is expressed at high levels within the pulmonary mesenchyme in a cell lineage-restricted pattern which is complementary to that of the GATA-6 gene. Finally, the GATA-5 gene is expressed in tissue-restricted subsets of SMCs, including bronchial SMCs and SMCs of the bladder wall. These data are consistent with a model wherein GATA-5 regulates a novel molecular program in the mammalian embryo and serves to identify GATA-5 as a candidate regulator of the transcriptional program(s) that underlies smooth muscle cell diversity.

Previous studies have suggested that GATA-4 may play

an important role in formation and/or specification of extra-embryonic tissues (Soudais *et al.*, 1995; Bielinska *et al.*, 1996). Soudais *et al.* (1995) reported that embryonic stem (ES) cells containing a homozygous null mutation in the GATA-4 gene exhibit gross defects in membranous structures including the visceral and parietal endoderm. While these data are consistent with the high levels of GATA-4 mRNA that were observed in these extra-embryonic tissues, the finding that the GATA-5 gene is not expressed either within the visceral or parietal endoderm suggests that the functional repertoire of GATA-5 is restricted in extra-embryonic tissues and limited to the allantois. Moreover, in the primitive streak embryo, the murine GATA-5 gene is expressed exclusively within the mesoderm subjacent to the headfold region which contains the cardiogenic plate. In contrast, the GATA-4 and -6 genes were developmentally coexpressed in the ED 7.0 embryo both within the primitive streak mesoderm as well as the cardiogenic plate. Once again, this suggests that the function of GATA-5 in the precardiac mesoderm is more restricted than that of either the GATA-4 or -6 genes. In this regard it is noteworthy that within the embryonic heart, the GATA-5 gene has a temporally and spatially restricted pattern of expression that is distinct from that of the GATA-4 and -6 genes (see below). Thus, identification of the subpopulation of cells within the lateral plate mesoderm that express GATA-5 may serve to define previously unrecognized sublineages within the primitive heart that give rise to atria versus ventricle and/or cardiac myocytes versus endocardial cells.

Several independent groups have suggested that the GATA-4/5/6 subfamily of transcription factors may play an important role in directing tissue-specific gene expression within the heart during embryonic and postnatal development (Jiang and Evans, 1996; Kelley *et al.*, 1993; Laverriere *et al.*, 1994; Grepin *et al.*, 1994; Ip *et al.*, 1994; Tamura *et al.*, 1993). In avian and *Xenopus* species the GATA-4, -5, and -6 genes are developmentally coexpressed within the precardiac mesoderm and embryonic and adult heart (Jiang and Evans, 1996; Laverriere *et al.*, 1994). Thus, it was somewhat surprising to find that in the mammalian species, the GATA-5 gene is expressed in a developmentally regulated, spatially restricted pattern that differs significantly from that of the mouse GATA-4 and -6 genes. Within the primitive heart tube, the GATA-5 gene is expressed at higher levels in the common atrial chamber than in the ventricle. In contrast, the GATA-4 and -6 genes are expressed at relatively equal levels in both the atria and ventricles (Morrisey *et al.*, 1996). Coincident with cardiac septation and formation of the endocardial cushions (ED 12.5), GATA-5 gene expression becomes restricted to the atrial endocardium, while the GATA-4 and -6 genes are expressed at high levels throughout the embryonic heart. Subsequently, during fetal and postnatal development, the GATA-5 gene is not expressed in the heart, whereas the GATA-4 and -6 genes continue to be expressed at high levels in both cardiac chambers (Heikinheimo *et al.*, 1994; Morrisey *et al.*, 1996). The unique pattern of GATA-5 gene expression within the mouse heart coupled with the demonstration that GATA-5 can transactivate a cardiac-specific transcriptional regulatory element in noncardiac muscle cells suggests that GATA-5 gene may play an important role in chamber-specification and/or establishment of the endocardium and/or endocardial cushions. These data support the emerging hypothesis that distinct transcriptional programs may direct atrial versus ventricular-specificity and left/right asymmetry within the vertebrate heart (Li *et al.*, 1996; Ross *et al.*, 1996; Tsuda *et al.*, 1996).

Fine differences in the patterns of expression of the GATA-1, -2, and -3 proteins in the hematopoietic cell lineages have translated into marked differences in the phenotypes of mice harboring null mutations in each of these family members (Pandolfi *et al.*, 1995; Pevny *et al.*, 1991; Simon *et al.*, 1992; Tsai *et al.*, 1994; Weiss *et al.*, 1994). Thus, the demonstration that the GATA-5 gene was expressed in the embryonic pulmonary mesenchyme and bronchial SMCs suggests that GATA-5 may mediate important, and unanticipated, functions in the embryonic lung. During expansion of the embryonic lung bud, the GATA-5 and -6 genes are expressed in complementary, cell lineage-restricted patterns; GATA-5 gene expression is restricted to the pulmonary mesenchyme and GATA-6 gene expression is restricted to the bronchial epithelial cells. Of note, the embryonic lung contains cells derived both from gut epithelium (bronchial epithelial cells), as well as lateral mesoderm (mesenchymal cells, endothelial cells, and smooth muscle cells). The finding that both the GATA-5 and -6 genes are expressed in the gut epithelium and lateral mesoderm, while these genes are differentially expressed in a cell lineage-specific pattern within the embryonic lung, suggests that GATA-5 and -6, in concert with other transcription factors, may play important roles in regulating the differentiation of distinct cell lineages during embryonic lung development.

Relatively little is currently understood about the molecular mechanisms that control SMC identity and differentiation (for review see Owens, 1995). Thus, the observation that GATA-5 is expressed in bronchial and bladder SMCs, but not in arterial or gut SMCs, suggests that GATA-5 may activate transcription in tissue-restricted subsets of the SMC lineage. Of note, we have reported recently that the GATA-6 gene is expressed within arterial, venous, and bladder SMCs (Morrisey *et al.*, 1996). Taken together, these data support and extend our hypothesis that distinct transcriptional programs may underlie SMC diversity and serve to identify GATA-5 (and GATA-6) as potential regulators of this program. As such, identification of the transcriptional targets of GATA-5 (and GATA-6) in the different SMC lineages should increase understanding of the molecular mechanisms that regulate SMC development and differentiation.

ACKNOWLEDGMENTS

We thank Jeff Leiden, Celeste Simon, and Beth McNally for reviewing the manuscript and providing helpful suggestions. We

thank Lisa Gottschalk for expert preparation of the figures and Amy Murphy for expert secretarial assistance. This work was supported in part by Public Health Service Grant R01HL51145 and an American Heart Association Grant-In-Aid to M.S.P. M.S.P. is an Established Investigator of the American Heart Association.

REFERENCES

Arceci, R. J., King, A. A., Simon, M. C., Orkin, S. H., and Wilson, D. B. (1993). Mouse GATA-4: A retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart. *Mol. Cell. Biol.* **13,** 2235–2246.

Bielinska, M., Narita, N., Heikinheimo, M., Porter, S. B., and Wilson, D. B. (1996). Erythropoiesis and vasculogenesis in embryoid bodies lacking visceral yolk sac endoderm. *Blood* **88,** 3720–3730.

Blobel, G. A., and Orkin, S. H. (1996). Estrogen-induced apoptosis by inhibition of erythroid transcription factor GATA-1. *Mol. Cell. Biol.* **16,** 1687–1694.

Crossley, M., Merika, M., and Orkin, S. H. (1995). Self-association of the erythroid transcription factor GATA-1 mediated by its zinc finger domains. *Mol. Cell. Biol.* **15,** 2448–2456.

Dorfman, D. M., Wilson, D. B., Bruns, G. A., and Orkin, S. H. (1992). Human transcription factor GATA-2. Evidence for regulation of preproendothelin-1 gene expression in endothelial cells. *J. Biol. Chem.* **267,** 1279–1285.

Evans, T., Reitman, M., and Felsenfeld, G. (1988). An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. *Proc. Natl. Acad. Sci. USA* **85,** 5976–5980.

Grepin, C., Dagnino, L., Robitaille, L., Haberstroh, L., Antakly, T., and Nemer, M. (1994). A hormone-encoding gene identifies a pathway for cardiac but not skeletal muscle gene transcription. *Mol. Cell. Biol.* **14,** 3115–3129.

Grepin, C., Robitaille, L., Antakly, T., and Nemer, M. (1995). Inhibition of transcription factor GATA-4 expression blocks in vitro cardiac muscle differentiation. *Mol. Cell. Biol.* **15,** 4095–4102.

Heikinheimo, M., Scandrett, J. M., and Wilson, D. B. (1994). Localization of transcription factor GATA-4 to regions of the mouse embryo involved in cardiac development. *Dev. Biol.* **164,** 361–373.

Ho, I. C., Vorhees, P., Marin, N., Oakley, B. K., Tsai, S. F., Orkin, S. H., and Leiden, J. M. (1991. Human GATA-3: a lineage-restricted transcription factor that regulates the expression of the T cell receptor alpha gene. *EMBO J.* **10,** 1187–1192.

Huang, X. (1994). On global sequence alignment. *Comput. Appl. Biosci.* **10,** 227–235.

Ip, H. S., Wilson, D. B., Heikinheimo, M., Tang, Z., Ting, C. N., Simon, M. C., Leiden, J. M., and Parmacek, M. S. (1994). The GATA-4 transcription factor transactivates the cardiac muscle-specific troponin C promoter-enhancer in nonmuscle cells. *Mol. Cell. Biol.* **14,** 7517–7526.

Jiang, Y., and Evans, T. (1996). The Xenopus GATA-4/5/6 genes are associated with cardiac specification and can regulate cardiac-specific transcription during embryogenesis. *Dev. Biol.* **174,** 258–270.

Joulin, V., Bories, D., Eleouet, J. F., Labastie, M. C., Chretien, S., Mattei, M. G., and Romeo, P. H. (1991). A T-cell-specific TCR delta DNA binding protein is a member of the human GATA family. *EMBO J.* **10,** 1809–1816.

Kelley, C., Blumberg, H., Zon, L. I., and Evans, T. (1993). GATA-4 is a novel transcription factor expressed in endocardium of the developing heart. *Development* **118,** 817–827.

Ko, L. J., and Engel, J. D. (1993). DNA-binding specificities of the GATA transcription factor family. *Mol. Cell. Biol.* **13,** 4011–4022.

Ko, L. J., Yamamoto, M., Leonard, M. W., George, K. M., Ting, P., and Engel, J. D. (1991). Murine and human T-lymphocyte GATA-3 factors mediate transcription through a cis-regulatory element within the human T-cell receptor delta gene enhancer. *Mol. Cell. Biol.* **11,** 2778–2784.

Kozak, M. (1984). Selection of initiation sites by eukaryotic ribosomes: Effects of inserting AUG triplets upstream from the coding sequence for preproinsulin. *Nucleic Acids Res.* **12,** 857–872.

Kuratani, S., Martin, J. F., Wawersik, S., Lilly, B., Eichele, G., and Olson, E. N. (1994). The expression pattern of the chick homeobox gene MHox suggests a role in patterning of the limbs and face and in compartmentalization of somites. *Dev. Biol.* **161,** 357–369.

Laverriere, A. C., MacNeill, C., Mueller, C., Poelmann, R. E., Burch, J. B., and Evans, T. (1994). GATA-4/5/6, a subfamily of three transcription factors transcribed in developing heart and gut. *J. Biol. Chem.* **269,** 23177–23184.

Lee, M.-E., Bloch, K. D., Clifford, J. A., and Quertomous, T. (1990). Functional analysis of the endothelin-1 gene promoter: evidence for an endothelial cell-specific cis-acting sequence. *J. Biol. Chem.* **266,** 16188–16192.

Li, L., Miano, J. M., Mercer, B., and Olson, E. N. (1996). Expression of the SM22α promoter in transgenic mice provides evidence for distinct transcriptional regulatory programs in vascular and visceral smooth muscle cells. *J. Cell Biol.* **132,** 849–859.

Merika, M., and Orkin, S. H. (1993). DNA-binding specificity of GATA family transcription factors. *Mol. Cell. Biol.* **13,** 3999–4010.

Molkentin, J. D., Kalvakolanu, D. V., and Markham, B. E. (1994). Transcription factor GATA-4 regulates cardiac muscle-specific expression of the α-myosin heavy chain gene. *Mol. Cell. Biol.* **14,** 4947–4957.

Morrisey, E. E., Ip, H. S., Lu, M. M., and Parmacek, M. S. (1996). GATA-6: A zinc finger transcription factor that is expressed in multiple cell lineages derived from lateral mesoderm. *Dev. Biol.* **177,** 309–322.

Morrisey, E. E., Ip, H. S., Tang, Z., and Parmacek, M. S. (submitted). GATA-4 activates transcription via two novel domains that are conserved within the GATA-4/5/6 subfamily.

Omichinski, J. G., Clore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J., and Gronenborn, A. M. (1993). NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science* **261,** 438–446.

Orkin, S. H. (1992). GATA-binding transcription factors in hematopoietic cells. *Blood* **80,** 575–581.

Osada, H., Grutz, G., Axelson, H., Forster, A., and Rabbitts, T. H. (1995). Association of erythroid transcription factors: Complexes involving the LIM protein RBTN2 and the zinc-finger protein GATA1. *Proc. Natl. Acad. Sci. USA* **92,** 9585–9589.

Owens, G. K. (1995). Regulation of differentiation of vascular smooth muscle cells. *Phys. Rev.* **75,** 487–517.

Pandolfi, P. P., Roth, M. E., Karis, A., Leonard, M. W., Dzierzak, E., Grosveld, F. G., Engel, J. D., and Lindenbaum, M. H. (1995). Targeted disruption of the GATA3 gene causes severe abnormalities in the nervous system and in fetal liver haematopoiesis. *Nature Genet.* **11,** 40–44.

Parmacek, M. S., Ip, H. S., Jung, F., Shen, T., Martin, J. F., Vora, A. J., Olson, E. N., and Leiden, J. M. (1994). A novel myogenic regulatory circuit controls slow/cardiac troponin C gene transcription in skeletal muscle. *Mol. Cell. Biol.* **14,** 1870–1885.

Parmacek, M. S., and Leiden, J. M. (1989). Structure and expression of the murine slow/cardiac troponin C gene. *J. Biol. Chem.* **264,** 13217–13225.

Parmacek, M. S., Vora, A. J., Shen, T., Barr, E., Jung, F., and Leiden, J. M. (1992). Identification and characterization of a cardiac-specific transcriptional regulatory element in the slow/cardiac troponin C gene. *Mol. Cell. Biol.* **12,** 1967–1976.

Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'Agati, V., Orkin, S. H., and Costantini, F. (1991). Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* **349,** 257–260.

Ross, R. S., Navankasattusas, S., Harvey, R. P., and Chien, K. R. (1996). An HF-1a/HF-1b/MEF-2 combinatorial element confers ventricular specificity and establishes an anterior-posterior gradient of expression. *Development* **122,** 1799–1809.

Simon, M. C. (1995). Gotta have GATA. *Nature Genet.* **11,** 9–11.

Simon, M. C., Pevny, L., Wiles, M. V., Keller, G., Costantini, F., and Orkin, S. H. (1992). Rescue of erythroid development in gene targeted GATA-1-mouse embryonic stem cells. *Nature Genet.* **1,** 92–98.

Solway, J., Seltzer, J., Samaha, F. F., Kim, S., Alger, L. E., Niu, Q., Morrisey, E. E., Ip, H. S., and Parmacek, M. S. (1995). Structure and expression of a smooth muscle cell-specific gene, SM22α. *J. Biol. Chem.* **270,** 13460–13469.

Soudais, C., Bielinska, M., Heikinheimo, M., MacArthur, C. A., Narita, N., Saffitz, J. E., Simon, M. C., Leiden, J. M., and Wilson, D. B. (1995). Targeted mutagenesis of the transcription factor GATA-4 gene in mouse embryonic stem cells disrupts visceral endoderm differentiation. *Development* **121,** 3877–3888.

Tamura, S., Wang, X.-H., Maeda, M., and Futai, M. (1993). Gastric DNA-binding proteins recognize upstream sequence motifs of parietal cell-specific genes. *Proc. Natl. Acad. Sci. USA* **90,** 10876–10880.

Tsai, F. Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F. W., and Orkin, S. H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature* **371,** 221–226.

Tsai, S. F., Martin, D. I. K., Zon, L. I., D'Andrea, A. D., Wong, G. G., and Orkin, S. H. (1989). Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature* **339,** 446–451.

Tsuda, T., Philp, N., Zile, M. H., and Linask, K. K. (1996). Left-right asymmetric localization of flectin in the extracellular matrix during heart looping. *Dev. Biol.* **173,** 39–50.

Weiss, M. J., Keller, G., and Orkin, S. H. (1994). Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. *Genes Dev.* **8,** 1184–1197.

Weiss, M. J., and Orkin, S. H. (1995a). GATA transcription factors: key regulators of hematopoiesis. *Exp. Hematol.* **23,** 99–107.

Weiss, M. J., and Orkin, S. H. (1995b). Transcription factor GATA-1 permits survival and maturation of erythroid precursors by preventing apoptosis. *Proc. Natl. Acad. Sci. USA* **92,** 9623–9627.

Wilson, D. B., Dorfman, D. M., and Orkin, S. H. (1990). A nonerythroid GATA-binding protein is required for function of the human preproendothelin-1 promoter in endothelial cells. *Mol. Cell. Biol.* **10,** 4854–4862.

Yamamoto, M., Ko, L. J., Leonard, M. W., Beug, H., Orkin, S. H., and Engel, J. D. (1990). Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. *Genes Dev.* **4,** 1650–1662.

Yang, Z., Gu, L., Romeo, P.-H., Bories, D., Motohashi, H., Yamamoto, M., and Engel, J. D. (1994). Human GATA-3 trans-activation, DNA-binding, and nuclear localization activities are organized into distinct structural domains. *Mol. Cell. Biol.* **14,** 2201–2212.

Zon, L. I., Gurish, M. F., Stevens, R. L., Mather, C., Reynolds, D. S., Austen, K. F., and Orkin, S. H. (1991). GATA-binding transcription factors in mast cells regulate the promoter of the mast cell carboxypeptidase A gene. *J. Biol. Chem.* **266,** 22948–22953.

Received for publication October 22, 1996
Accepted November 25, 1996